Multi-Device Selection Scheduling in Non-Identically Distributed Fading Channels

Ko, Y., Quddus, A. U. I., & Tafazolli, R. (2016). Multi-Device Selection Scheduling in Non-Identically Distributed Fading Channels. IET Communications. DOI: 10.1049/iet-com.2015.0977

Published in:
IET Communications

Document Version:
Peer reviewed version

Multi-Device Selection Scheduling in Non-Identically Distributed Fading Channels

Youngwook Ko, *Member, IEEE,* Atta Ul Quddus, and Rahim Tafazolli *Senior Member, IEEE*

Abstract

Multiuser selection scheduling concept has been recently proposed in the literature in order to increase the multiuser diversity gain and overcome the significant feedback requirements for the opportunistic scheduling schemes. The main idea is that reducing the feedback overhead saves per-user power that could potentially be added for the data transmission. In this work, we propose to integrate the principle of multiuser selection and the proportional fair scheduling scheme. This is aimed especially at power-limited, multi-device systems in non-identically distributed fading channels. For the performance analysis, we derive closed-form expressions for the outage probabilities and the average system rate of the delay-sensitive and the delay-tolerant systems, respectively, and compare them with the full feedback multiuser diversity schemes. The discrete rate region is analytically presented, where the maximum average system rate can be obtained by properly choosing the number of partial devices. We optimize jointly the number of partial devices and the per-device power saving in order to maximize the average system rate under the power requirement. Through our results, we finally demonstrate that the proposed scheme leveraging the saved feedback power to add for the data transmission can outperform the full feedback multiuser diversity, in non-identical Rayleigh fading of devices' channels.

Index Terms

Opportunistic scheduling, reduced feedback, multiuser selection scheduling, heterogeneous fading

I. Introduction

In wireless communications, the significant increase of the number of battery-powered devices over the last decade has drawn research interest in the area of power-efficiently scheduling them

Youngwook Ko is with the Institute of Electronics, Communications and Information Technology, Queen's University Belfast, BT3 9DT, UK; and Atta Ul Quddus and Rahim Tafazolli are with the Institute for Communication Systems in the Department of Electronic Engineering, University of Surrey, Guildford, Surrey, GU2 7XH, UK, e-mails: (y.ko@qub.ac.uk, {a.quddus, r.tafazolli}@surrey.ac.uk).

to the shared wireless resources, for example, 1000 wireless devices per person are envisaged by 2020 [1]. With such large number of distributed devices, the concept of multi-user scheduling scheme, which use independent and time-varying multipath fading of users' channels for exploiting multiuser diversity (MUDiv), is important particularly for power-constrained wireless systems, such as sensor systems.

The MUDiv has been well studied in the literature [2]–[5] and is included in the next evolution of WiFi (e.g., IEEE 802.11 ac). Exploiting the MUDiv gain, advanced scheduling schemes have been extensively developed in [4], [5]. In the MUDiv systems, the random channel fading condition is treated as an opportunistic resource. That is, the user having the most favorable channel fading condition is opportunistically scheduled to transmit/receive over the entire transmission interval. For example, in maximum-rate scheduling (referred to as greedy multiuser scheduling) [4], [5], the "best" user having the maximum signal-to-noise ratio (SNR) is scheduled and thus the system rate is maximized. However, users suffering from poor channel conditions (due to, e.g., highly shadow fading) may be deprived from gaining access to the channel. To avoid such disadvantage, maximum normalized-SNR scheduling (referred to as proportional-fair multiuser scheduling) (see [5] and references therein) schedules the user having the maximum SNR (normalized to its own average gain). This way of scheduling maximizes the system rate while guaranteeing the rate of each user proportional to the user's channel condition, and provides proportional fairness. The MUDiv gain is applied to a range of emerging multiuser applications, e.g., multiuser spectrum aggregation, simultaneous wireless information and power transfer (SWIPT), multiuser multiple-input multiple-output (MU-MIMO) systems, orthogonal frequency-division multiple access (OFDMA) systems and the third generation partnership programme (3GPP)'s small cells [6]–[12].

To further improve the MUDiv gain, practical scheduling strategies and various system performance measures have been investigated [5], [13]–[16]. For example, the performance of the conventional MUDiv is limited as the required channel-state-information feedbacks are too complex to operate in a time-varying channel. In [13], it is shown that, due to the fact that the feedback channels are in practice outdated, determining the best user over the user scheduling steps cannot be realistically computed and this hinders the use of any effective scheduling. In [14], the sum capacity imposed by the MUDiv has been studied with respect to two MUDiv system performance measures such as scheduling complexity and scheduling fairness. In [15],

partial users and limited bits feedback were studied to reduce the feedback overhead. For small cell multiuser systems, a joint admission and power control method was studied to maximize the number of acceptable users in [16]. In [5], it is argued that, to overcome the heavy feedback requirements of the conventional scheduling schemes, reducing the achievable MUDiv gain can be an acceptable traded-off for saving the required channel feedbacks.

As for the reduced number of required channel feedbacks, several MUDiv methods have been recently presented in [17]–[20]. The idea is to allow only a fraction of the users to be active in the MUDiv schemes, being not limited to full channel feedbacks (e.g., full feedback multiuser scheduling). In [17], the authors proposed the multiuser switched-diversity scheduling scheme, where the idea is to find any acceptable user under good channel condition, instead of finding the best user among all. The enhanced multiuser switched-diversity scheduling scheme is presented in [18], where the concept of the per-user thresholds was suggested to improved the performance of the multiuser switched-diversity scheduling scheme. On the other hand, the authors in [19] proposed multiuser selective scheduling scheme in which a fraction of random users send the required feedbacks and per-user power saving obtained by the reduced feedbacks is added for the data transmission. Considering independent, and identically distributed fading of users' channels, [19] showed that decreasing the number of active users (i.e., required feedbacks) can decrease the overall bit error probability along with the reduced channel feedbacks. In [20], two types of channel feedback methods were proposed, both quantifying how many users should feedback channel information, i.e., the amount of the available MUDiv to be used from the perspective of the system throughput. However, in the existing literature, the MUDiv systems have been investigated with fixed power allocation to the data: reducing the required feedbacks does not influence the potential data power amount. To the best of the authors' knowledge, multi-user scheduling schemes that investigate *jointly the reduced feedbacks and opportunities of the associated power saving per user* have not been analyzed, especially focusing on *heterogeneous fading multi-device systems*, which differs from [19] addressing only homogeneously fading cases. Taking into account the increasing number of the heterogeneous fading user devices in the future wireless systems, it is important to study the sensitivity of the MUDiv system to the heterogeneous fading under the limited power usage.

In this work, we consider power-efficient multi-device scheduling scheme over non-identically distributed fading channels. Particulary, extending the work in [19], we propose proportional-

fair multiuser selective scheduling scheme over non-identically distributed fading multi-device systems. Consider two cases: (1) delay-sensitive case; (2) delay-tolerant case. Thus, this work aims not only to evaluate outage probability, but also to maximize system rate under the given power requirements. To this end, we mathematically analyze both the cumulative distribution functions and the average system rate, deriving their closed-form expressions over the heterogeneous Rayleigh fading of devices' channels. Towards improving the performance, we optimally develop the opportunity for jointly finding the number of active devices and exploiting the per-device potential power saving by reduced feedbacks. Based on our results, it will be demonstrated that compared to the conventional schemes, introducing only a subset of devices chosen randomly to the scheduling is a better choice not only to increase the average system rate, but also to decrease the outage probability. Higher the heterogeneous fading, larger the achievable MUDiv gain is. Referring to the optimum results, it will be demonstrated that a percentage of the active devices among the available ones should decrease in order to maximize the average system rate when the number of available devices increases.

The paper is organized as follows. The system and channel models are introduced in Section II, and the proportional fair multiuser selection scheduling scheme is described formally in Section III. In Section IV, the power-loading balance and the outage probability of the proposed scheme are addressed, followed by the mathematical analysis of the sensitivity of the proposed scheme in the non-identically distributed fading case. Section V contains the analysis of the average system rate and, in Section VI, the extension to multiple antenna systems is discussed, deriving upper bound expressions for the average system rate and comparing it to the conventional schemes. In Section VII, the optimization solution to the maximum average system rate is presented, while in Section VIII further asymptotic analysis for several cases are provided. Section IX presents numerical and simulation results and is followed by conclusions in Section X.

II. System model

A. System model

We now outline the system model for our power-limited multi-device scheduling scheme. Suppose that N distributed devices opportunistically communicate to the central device unit (CU). For the uplink transmission by N devices, the signal to be transmitted is denoted by s_l

for $l \in \{1, \cdots, N\}$. The received signal at the CU with M receiving antennas is given by

$$\mathbf{r}_l = \mathbf{h}_l s_l + \mathbf{n}_l \tag{1}$$

where $\mathbf{n}_l$ denotes the complex-valued zero-mean additive background noise vector, i.e., $\mathbf{n}_l \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$, and $\mathbf{h}_l = \sqrt{\rho_l}\,\mathbf{g}_l = \left[\sqrt{\rho_l}\,g_{l,0}, \cdots, \sqrt{\rho_l}\,g_{l,M-1}\right]^T$ is the general composite fading channel vector where its element $\sqrt{\rho_l}\,g_{l,m}$ represent the Rayleigh fading channel, at receiving antenna m, being independent and non-identically distributed complex Gaussian, i.e., $\mathbf{h}_l \sim \mathcal{CN}(\mathbf{0}, \rho_l \mathbf{I})$, and $\rho_l \neq \rho_j, \forall l, j$. Due to dynamic environments, here, ρ_l are modelled as random variables resulting from both user-specific shadowing and path loss.

In the conventional device-to-device system, each data symbol is an element of a given constellation. That is, $s_l \in \mathcal{S}$ where $\mathcal{S}$ represents the signal constellation. Suppose that $\mathbf{h}_l$ is known at the CU, the data rate of device l per channel use can be obtained as

$$R_l = \log_2\left(1 + P_{\underline{d}} \|\mathbf{h}_l\|^2\right),$$

where $P_{\underline{d}}$ denotes the transmission power of s_l and the normalized noise variance in (1) is used. Thus, the performance depends on the channel vector, $\mathbf{h}_l$, and no scheduling diversity can be exploited.

B. Selective Scheduling for multi-device MIMO

Denote by l^* the index of the device scheduled for the data transmission. Then, the received signal from device l^* is

$$\mathbf{r}_{l^*} = \mathbf{h}_{l^*} s_{l^*} + \mathbf{n}_{l^*} \tag{2}$$

where $\mathbf{h}_{l^*} \in \Sigma_h = \{\mathbf{h}_{(1)}, \cdots, \mathbf{h}_{(n)}\}$ and $\Sigma_h \subseteq \{\mathbf{h}_1, \cdots, \mathbf{h}_N\}$. In selective scheduling (SS), only n devices are chosen to participate in scheduling and contribute to data symbols. To this end, a subset Σ_h and its size n are properly chosen every scheduling interval, in the following.

Denote by Σ_D a subset of n devices. For a given N, let the CU properly choose $n(\leq N)$ and comprises of a Σ_D with no priori knowledge of the channel state information (CSI). The subset is given by

$$\Sigma_D = \{D_{(1)}, \cdots, D_{(n)}\} \tag{3}$$

where $D_{(l)}$ denotes the (l)th device element of Σ_D and $\Sigma_D \subseteq \{D_1, \cdots, D_N\}$. For example, given $N = 10$ devices and $n = 4$, the CU receives s_{l^*} opportunistically from only 4 among

10 devices. It differs from the conventional scheduling schemes, e.g., proportional fair (PF) scheduling, which receives opportunistically from all N devices.

Provided Σ_D is chosen, the SS applies the principle of the PF scheduling to $D_{(l)}$'s. Particularly, the SS has following subsequent planes: (1) multiple access control (MAC) plane; and (2) multi-device diversity (MD) plane.

The MAC plane randomly comprises of Σ_D in (3) and activates the subset of n devices in Σ_D, every scheduling interval. It requires one feedback bit to each device for the acknowledgement of its activation. Here, let $D_l, \forall l$ be equally likely chosen as an element of Σ_D, i.e., mathematically D_l has the uniform probability n/N of its activation. Given $D_{(l)}$'s in Σ_D, then, a combination of n sub-channels is reserved for the training transmission between the CU and $D_{(l)}$'s. The CU is assumed to obtain the CSI, Σ_h, of the $D_{(l)}$'s.

The MD plane refers to Σ_h in which CU schedules the best device to the best sub-channel for the data transmission. The scheduling criterion is to find the best among $D_{(l)}$'s having the relatively best channel to its own channel statistics. Thus, l^* in (2) can be given by

$$l^* \triangleq \arg\max_{(l)} \frac{\|\mathbf{h}_{(l)}\|^2}{\rho_{(l)}}. \tag{4}$$

Here, employing the principle of the PF scheduling, the relatively best channel for the scheduling is considered. This ensures that each device can be scheduled at an equal probability, due to the fact that $\mathbf{h}_{(l)}/\sqrt{\rho_{(l)}} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$, and detailed proof is referred to Appendix. Based on (4), the finite n bits feedback (i.e., one feedback bit to $D_{(l)}$) are employed to inform $D_{(l)}$'s a binary decision on, that is, accessing (or not accessing) the channel. The data rate by the SS is given by

$$R = \log_2\left(1 + \rho_{l^*}\|\mathbf{g}_{l^*}\|^2 P_d\right) \tag{5}$$

where P_d denotes the data power by the SS and recall that $\mathbf{g}_a = \mathbf{h}_a/\sqrt{\rho_a}$, being the normalized channel vector, i.e., $\mathbf{g}_a \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$, at device a.

III. Power-Loading Balance and Performance Analysis

We now analyze the performance of the proposed scheme for two cases: (i) delay-sensitive multi-device system; and (ii) delay-tolerant multi-device system. We examine the outage probability and the average system rate for cases (i) and (ii), respectively. We assume that $\mathbf{h}_l$ are

non-identically distributed fading with their shadowing $\rho_l \neq \rho_j, \forall l, j$. We first address power-loading balance for the scheduling.

A. Training and data power-loading balance

Denote by P_T the total transmit power by the n devices. The transmit power of both the training and the data by the n devices can be balanced so that every realization of Σ_h, the instantaneous power usages remain below the target level, P_0. We will refer to this balance as "power-loading balance". It is also worth mentioning that the power-loading balance is associated with a feedback cost. That is, a feedback cost of the multiuser scheduling can be quantified with the number of feedback users (e.g., see [5], [20]) or with the sum of per-user feedback message power (or training power).

In the conventional scheduling schemes, N devices transmit the training and the best among them transmits the data. Thus, P_T imposed by the N devices can be

$$P_T = N P_{\underline{d}}\, \alpha + P_{\underline{d}} \tag{6}$$

where α denotes the given power ratio of the training and the data transmissions. That is, $(P_{\underline{d}}\, \alpha)$ is used to quantify the training power (as feedback cost) by each device, while $P_{\underline{d}}$ is for the data power.

On the other hand, the SS enables a subset of n devices to transmit the training and the data. The power of the training (as feedback cost) by n devices can be small for a given N. Thus, P_T of the SS is

$$P_T = n P_{\underline{d}}\, \alpha + P_d(n). \tag{7}$$

Inserting (6) into (7), the data power of the SS can be chosen as a function of n:

$$P_d(n) = (N\alpha + 1)\, P_{\underline{d}} \quad - \underbrace{n\, P_{\underline{d}}\, \alpha}_{\text{Training power}} \tag{8}$$

$$= P_{\underline{d}} + \underbrace{(N-n)\, P_{\underline{d}}\, \alpha}_{\text{Potential data power}} \tag{9}$$

In (8), $P_d(n)$ can be obtained by subtracting from P_T in (6) the training power by the n devices. Smaller n, larger the remaining power $(N-n)P_{\underline{d}}\, \alpha$ is saved. This saving power could be potentially added for $P_d(n)$ in (9).

It is worth pointing out that unlike the conventional schemes, the SS can take into account the above power-loading balance between the training and the data power through the ratio n/N. This flexibility will be leveraged to improve the average system rate later in this paper.

B. Outage Probability

Consider a delay-sensitive multi-device case. As a measure of merit, we analyze the outage probability, deriving its expression with the statistics of the best SNR over the heterogeneous shadowing environments.

Firstly denote by $F_y(\cdot)$ the cumulative distribution function (cdf) of y. When $y = \rho_{l^*}\|\mathbf{g}_{l^*}\|^2$ in (5), $F_y(x)$ is also referred to as the outage probability that the channel strength of the best device l^* remains below x, and is given by

$$F_y(x) = \Pr\left[\rho_{l^*}\|\mathbf{g}_{l^*}\|^2 \leq x\right]. \tag{10}$$

We rewrite (10) for further analysis. Particularly, notice the fact that the fairness in scheduling is obtained among devices via subset Σ_D, since the best index l^* is selected according to its potential channel gain related with the average gain. For given n and Σ_D, thus, $D_{(l)}$ in Σ_D are uniformly chosen as the best with the equal probability of $1/n$. Also, referring to (4), the best index l^* relies on $\|\mathbf{g}_{l^*}\|^2$, statistically being independent of $\rho_{(l)}$. Based on this observation, $F_y(x)$ in (10) can be re-written, using the union bound expression [21], as

$$F_y(x) \leq \frac{1}{n}\sum_{(l)=1}^{n}\Pr\left[\rho_{(l)}\|\mathbf{g}_{l^*}\|^2 \leq x\right] \tag{11}$$

$$= \frac{1}{n}\sum_{(l)=1}^{n}\Pr\left[\|\mathbf{g}_{l^*}\|^2 \leq \frac{x}{\rho_{(l)}}\right] \tag{12}$$

where $\Pr[\cdot]$ in (12) represents the cdf of the normalized best channel gain. Using the higher order statistics [22], this probability can be obtained as

$$\Pr\left[\|g_{l^*}\|^2 \leq \frac{x}{\rho_{(l)}}\right] = \gamma\left(1, \frac{x}{\rho_{(l)}}\right)^n$$

where $\gamma(1,z) \triangleq (1-e^{-z})$, when $M = 1$ at the CU. Hereinafter, for simple analysis, $M = 1$ is considered and its extension to multiple antenna system will be addressed in Section IV.

Notice that the inequality in (11) can become the equality in the SS case due to the fact that only one device as the best is scheduled every channel use. Using (12), therefore, we can express

(10) in the following form, for given Σ_D and n, as

$$F_y(x) = \frac{1}{n} \sum_{(l)=1}^{n} \gamma \left(1, \frac{x}{\rho_{(l)}}\right)^n. \tag{13}$$

To validate the accuracy of (13), Fig. 1 depicts the comparison of the simulated cdfs of $y = \rho_{l^*} \|g_{l^*}\|^2$ with the theoretical cdfs of (13) for various n's. For the illustrations on this purpose, we use when $n = N \in \{5, 10, 20, 50\}$ devices and $\rho_{(l)}$ are heterogeneous, log-normal shadowing with standard deviation σ, i.e., $\sigma = 1$ (dB). As seen in Fig. 1, (13) is very close to the simulation results for various n's, which validates the accuracy of (13).

Also, it is worth mentioning from Fig. 1 that with larger n, $F_y(\cdot)$ increases slower. This indicates that the heterogeneous multi-device system is less influenced by the channel hardening effect [23] predominant in homogeneous case. Particularly, when $n = 5$ increases to $n = 50$, $F_y(\cdot) = 0.6$ can be reduced down to 0.08. This implies that, for given $n = 50$ devices, 55 per-cent less outage can occur.

Fig. 2 illustrates the impact of the heterogenity on the theoretical cdfs in (13). To validate its accuracy, (13) is compared with the simulated cdfs for various σ's. As seen in the figure, (13) performs close to the simulations. Also, it can be seen in the figure that as σ decreases, the achievable cdfs are closer to each other. This leads to the fact that less heterogeneous fading results in higher outage probability. In other words, the more the heterogeneous fading, the larger the scheduling power gain.

C. Average system rate

Consider a delay-tolerant multi-device case. According to the information theory, one can achieve the channel capacity through an extremely large number of coding bits. Using such large number of coding bits, the achieved channel capacity is available with long delay and is valuable as performance metric, particularly in a delay tolerant system. The channel capacity can be referred to as the achievable system rate in our work. Accordingly, the system rate of our work is achieved at long delay and thus, the corresponding system can be referred to as a delay-tolerant system For this, denote by $\bar{R}$ the average system rate. Based on the results in Section III, $\bar{R}$ can be formulated as the weighted sum of the average individual rates. In

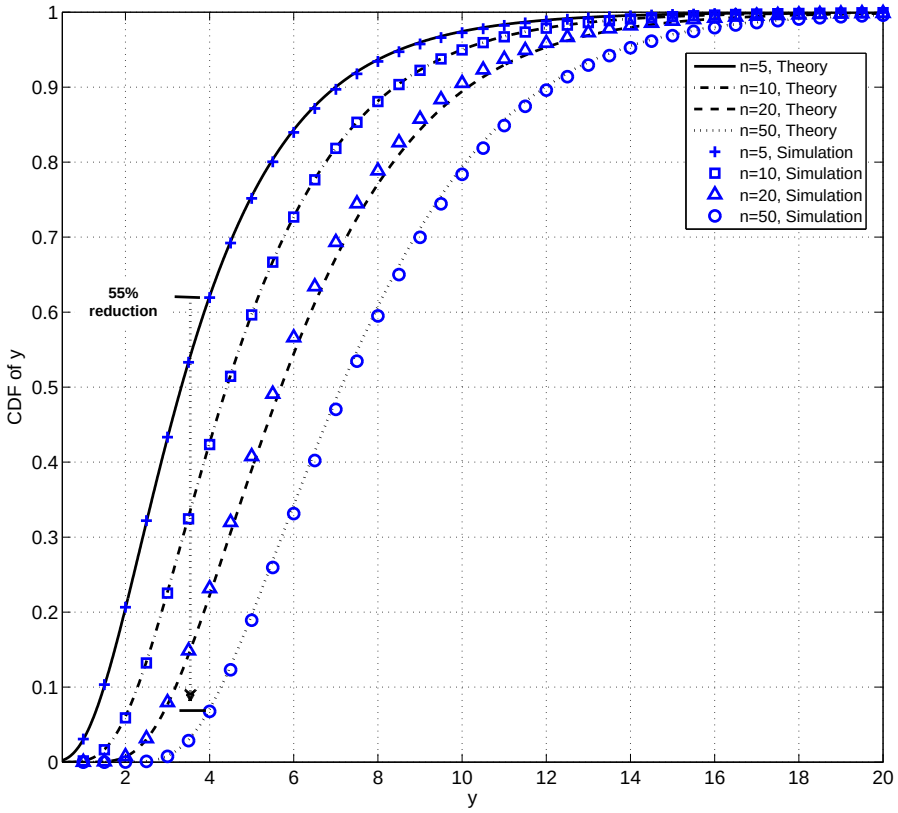

Fig. 1. Cumulative distribution functions of a random variable $y = \rho_{l^*} \|\mathbf{g}_{l^*}\|^2$ where the theoretical value in (13) is shown without markers and the simulated y are shown with makers only, and $n = \{5, 10, 20, 50\}$.

particular, when $n \leq N$, we have, using (5) and (13),

$$\bar{R} = \frac{1}{N} \sum_{l=1}^{N} \int_0^\infty \log_2 \left(1 + \rho_l P_d(n)\, x\right) p_{\|g_{l^*}\|^2}(x)\, dx \tag{14}$$

where $p_{\|g_{l^*}\|^2}(\cdot)$ denotes the probability density function (pdf) of $\|g_{l^*}\|^2$. The equality is based on the fact observed in (13): the best index l^* is related to $\|g_{l^*}\|$, being independent of ρ_l.

Applying the higher order statistics [22], when $M = 1$, $p_{\|g_{l^*}\|^2}(\cdot)$ in (14) is found, for given N and n, as

$$p_{\|g_{l^*}\|^2}(x) = n\, e^{-x} \gamma(1, x)^{n-1}.$$

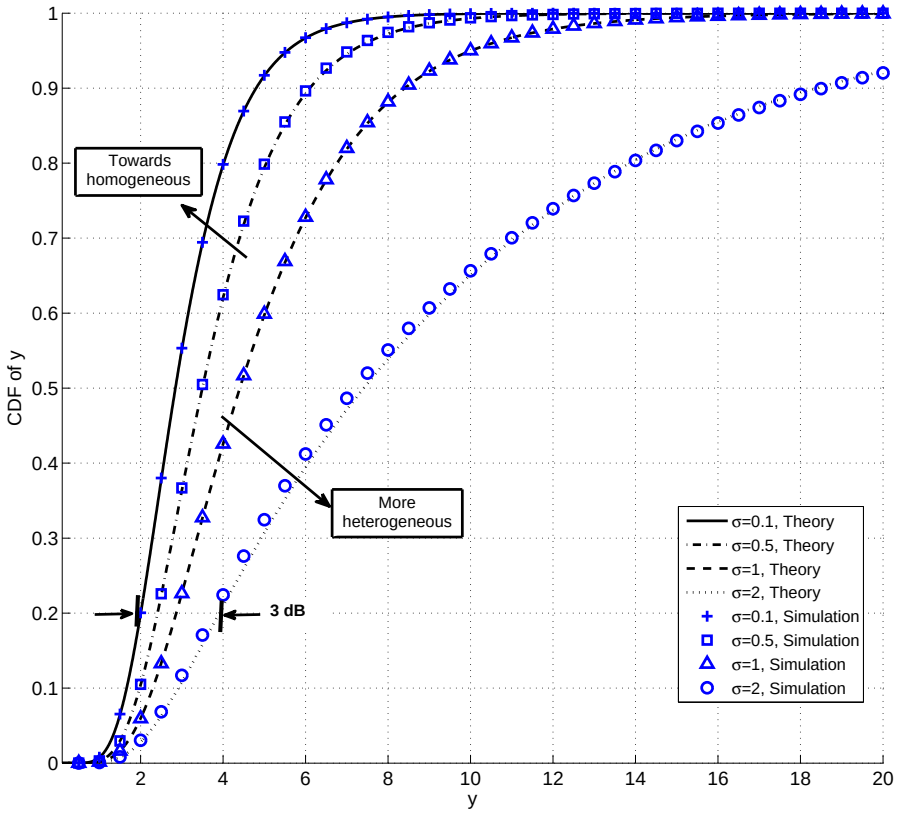

Fig. 2. Cumulative distribution functions of a random variable y for various heterogeneous fading when $y = \rho_{l^*} \|\mathbf{g}_{l^*}\|^2$, $n = 10$, and the heterogeneous fading components $\sigma \in \{0.1, 0.5, 1, 2\}$ (dB).

Inserting $p_{\|g_{l^*}\|^2}(\cdot)$ and $P_d(n)$ in (9) into (14), thus, $\bar{R}$ can be written as

$$\bar{R} = \frac{n}{N} \sum_{l=1}^{N} \int_0^{\infty} \log_2 \left(1 + P_d(n)\, \rho_l\, x\right) e^{-x} \gamma(1, x)^{n-1}\, dx \tag{15}$$

where we have $P_d(n) = ((N - n)\, \alpha + 1)\, P_{\underline{d}}$ as a decreasing function of n. It can be seen in (15) that $\bar{R}$ increases with a proper selection of n, for given N and $P_{\underline{d}}$.

Refer to the fact that $\int_0^{\infty} \log_2(1+ax)p(x)dx = a/\ln 2 \int_0^{\infty}(1-F(x))/(1+ax)dx$, where $p(x)$ and $F(x)$ are the pdf and the cdf of x, respectively. Using this fact, the integral in (15) can be given by

$$\frac{a}{\ln 2} \int_0^{\infty} \frac{1 - F(x)}{1 + ax} dx,$$

where $F(x) = (1 - \exp(-x))^n$ and $a = P_d(n)\rho_l$. Using the Taylor series formulation and [24, (3.352.4)] and after simplifications, this integral can be given in closed-form as

$$\frac{a}{\ln 2}\int_0^\infty \frac{1-F(x)}{1+ax}dx = \frac{-1}{\ln 2}\sum_{m=1}^{n}\binom{n}{m}(-1)^{m+1}\,\mathrm{e}^{m/a}\,\mathrm{Ei}\left(-\frac{m}{a}\right),$$

where $\mathrm{Ei}(\cdot)$ is the exponential integral function. From these, the exact expression for (15) can be obtained in closed-form as

$$\bar{R} = \frac{-n}{N\ln 2}\sum_{l=1}^{N}\sum_{m=1}^{n}\binom{n}{m}(-1)^{m+1}\,\mathrm{e}^{m/a}\,\mathrm{Ei}\left(-\frac{m}{a}\right), \tag{16}$$

where recall that $a = P_d(n)\rho_l$.

IV. Extension to multiple antenna multi-device system

The performance analysis in the previous section can be extended to multiple antenna multi-device systems. This will also allow to use the benefits of multiple antenna techniques. Toward this end, a generalized expression for the average system rate has to be derived. For brevity, we address only the approximate average system rate case.

A. Upper bound expression for the average system rate

Let M be the multiple antenna diversity order. Then, in the multiple antenna case, the degrees of freedom (DOF) of x in (15) extends to $2M$, that is, $x \sim \chi^2_{2M}$ where χ^2_{2M} stands for the Chi-squared distribution with $2M$ DOF (refer also to [19]). Therefore, for given $n(\leq N)$ and M, the expression for $p_{\|\mathbf{g}_{l^*}\|^2}(x)$ in (14) can be generalized as [3]

$$p_{\|\mathbf{g}_{l^*}\|^2}(x) = n\frac{(x)^{M-1}\,\mathrm{e}^{-x}}{\Gamma(M)^n}\gamma\left(M,x\right)^{n-1} \tag{17}$$

where $\gamma(\cdot,\cdot)$ and $\Gamma(\cdot)$ represent the incomplete gamma and the complete gamma functions, respectively [24].

Inserting (17) into (14), $\bar{R}$ of the multiple antenna case can be

$$\bar{R} = \frac{1}{N}\sum_{l=1}^{N}\mathbb{E}\left(\log_2\left(1+\rho_l\|\mathbf{g}_{l^*}\|^2 P_d(n)\right)\right)$$

where $\mathbb{E}(\cdot)$ stands for the expectation operator. Using Jensen's inequality, we can have

$$\bar{R} \leq \frac{1}{N}\sum_{l=1}^{N}\log_2\left(1+\rho_l\,P_d(n)\,\mathbb{E}\|\mathbf{g}_{l^*}\|^2\right). \tag{18}$$

To further analysis, we refer to the fact that when n is large, the distribution of $\|\mathbf{g}_{l^*}\|^2$ in (18) satisfies [25]:

$$\begin{aligned} \left(\|\mathbf{g}_{l^*}\|^2 - \mu_{(n,M)}\right)/\beta_{(n,M)} &= z, \\ z \sim \mathcal{G}(z) &= \exp\left(-\exp(-z)\right) \end{aligned} \tag{19}$$

where $\mu_{(n,M)} = F_x^{-1}(1-1/n)$, $\beta_{(n,M)} = F_x^{-1}(1-1/n\mathrm{e}) - \mu_{(n,M)}$, $F_x(\cdot) = \gamma(M,x)/\Gamma(M)$ is the cdf of $\|\mathbf{g}_{(l)}\|^2$, z denotes a Gumbel distribution random variable, and $\mathcal{G}(z)$ denotes the cdf of z.

Using this asymptotic distribution (19), (18) can be approximated in closed–form as

$$\bar{R} \cong \frac{1}{N}\sum_{l=1}^{N} \log_2\left(1 + \rho_l P_d(n)\left(\mu_{(n,M)} + \beta_{(n,M)}\omega\right)\right) \tag{20}$$

where ω denotes the mean of z (i.e., $\mathbb{E}(z)$) that is Euler's constant.

B. Comparison with the conventional scheme

1) Power gain for the same average system rate: Denote by $\bar{R}_c$ the average system rate of the conventional scheme where $n = N$ is fixed. Accordingly, $\bar{R}_c$ is a special case of $\bar{R}$ with $n = N$ and we have

$$\bar{R}_c = \frac{1}{N}\sum_{l=1}^{N} \log_2\left(1 + \rho_l P_{\underline{d}}\left(\mu_{(N,M)} + \beta_{(N,M)}\omega\right)\right).$$

For comparison to $\bar{R}_c$, $\bar{R}$ can be represented with respect to only n and $P_{\underline{d}}$, for a given $n \leq N$:

$$\bar{R} \cong \frac{1}{N}\sum_{l=1}^{N} \log_2\left(1 + \rho_l\, P_{\underline{d}}\, G(n)\left(\mu_{(N,M)} + \beta_{(N,M)}\omega\right)\right) \tag{21}$$

where $G(n)$ is given by

$$G(n) = ((N-n)\alpha + 1)\,\frac{\mu_{(n,M)} + \beta_{(n,M)}\omega}{\mu_{(N,M)} + \beta_{(N,M)}\omega} \tag{22}$$

and denotes the *achievable power gain*. Notice that $G(n)$ is achieved, using the power-loading balance under the total power requirement ($P_T \leq P_0$) and the selective multi-device scheduling.

It can be shown from (21)–(22) that, for given N and ρ, $\bar{R}$ *benefits from properly selecting* n. In particular, for given N and ρ_l, $\bar{R}$ depends only on the achievable gain $G(n)$. From $G(n)$ in (22), notice that, as n increases, the terms $((N-n)\alpha + 1)$ and $\left(\mu_{(n,M)} + \beta_{(n,M)}\omega\right)$ monotonically decrease and increase, respectively. The latter is due to the fact that $\mu_{(n,M)}$ and $\beta_{(n,M)}$ are monotonic in terms of n, according to their definitions in (19). The former results from the power-loading balance through n. Therefore, it is inferred from (21)–(22) that *a proper selection of* n *can*

increase $G(n)$, taking into account the trade-off between $(N\nu - n\nu + \alpha)$ and $\left(\mu_{(n,M)} + \beta_{(n,M)}\omega\right)$, and thus, higher the average system rate $\bar{R}$.

2) Power gain for the same outage probability: Denote by $F_{y_c}(\cdot)$ the cdf of y_c and by y_c the normalized SNR, i.e., $y_c = \rho_{l^*}\|\mathbf{g}_{l^*}\|^2$. Conventionally let the data power $P_d(N)$ be fixed for all n's. Accordingly, we have, for a given threshold μ,

$$\begin{aligned}\Pr\left[y_c\, P_d(N) \le \mu\right] = F_{y_c}\left(\frac{\mu}{P_d(N)}\right) &= \frac{1}{N}\sum_{l=1}^{N}\Pr\left[\|\mathbf{g}_{l^*}\|^2 \le \frac{\mu}{\rho_l P_d(N)}\right] \\ &= \frac{1}{N}\sum_{l=1}^{N} F_x\left(\frac{\mu}{\rho_l P_d(N)}\right)^n,\end{aligned} \tag{23}$$

where recall that $F_x(x) = \gamma(M,x)/\Gamma(M)$.

For comparison to $F_{y_c}(\cdot)$, the cdf of the proposed scheme $F_y(\cdot)$ can be given with a power gain, for a given $n \le N$, as:

$$\begin{aligned}F_y\left(\frac{\mu}{P_d(n)}\right) &= \frac{1}{N}\sum_{l=1}^{N} F_x\left(\frac{\mu}{\rho_l P_d(n)}\right)^n \\ &= \frac{1}{N}\sum_{l=1}^{N} F_x\left(\frac{\mu}{\rho_l P_d(N)}\,\frac{1}{G_o(n)}\right)^n \\ &= F_{y_c}\left(\frac{\mu}{P_d(N)G_o(n)}\right),\end{aligned} \tag{24}$$

where $G_o(n) = 1 + (N-n)\alpha$.

From (23)-(24), it can be clearly found that the power gain achievable for the same outage probability is $G_o(n) = 1 + (N-n)\alpha$ and decreases monotonically with n.

V. Optimization problem

We formulate the optimization problem in order to maximize $\bar{R}$ under the total power requirement at the devices level. In particular, notice from (21)–(22) that $\bar{R}$ can be given as a function of only n for given N and M. Thus, for given N and M, the problem is posed as

$$\begin{aligned}\max_{n\in\{1,\cdots,N\}} \quad \bar{R} = \max_{n\in\{1,\cdots,N\}} \frac{1}{N}\sum_{l=1}^{N}\log_2\left(1 + \rho_l\, \underline{P_d}\, G(n)\left(\mu_{(N,M)} + \beta_{(N,M)}\omega\right)\right) \\ \text{subject to} \quad P_T \le P_0.\end{aligned} \tag{25}$$

Interestingly, notice from (25) that $\bar{R}$ can increase with the achievable power gain $G(n)$, regardless of ρ_l's. This observation leads to the fact that the value n maximizing $G(n)$ is

eventually maximizing $\bar{R}$, resulting in the variable power ratio between the training and the data in (7). Therefore, we equivalently present the following objective, maximizing $G(n)$ for given N and M, as

$$\max_{n} G(n) = \max_{n\in\{1,\cdots,N\}} \left((N-n)\alpha+1\right) \frac{\mu_{(n,M)}+\beta_{(n,M)}\omega}{\mu_{(N,M)}+\beta_{(N,M)}\omega} \quad \text{subject to} \quad P_T \leq P_0. \tag{26}$$

To solve this problem, let us consider a simple example, having a single receiving antenna $M=1$. In this context, we can obtain that $\mu_{(N,1)}=\log(N)$ and $\beta_{(N,1)}=1$, straightforwardly from their definitions in Section IV-A. Using these, it can be shown that (26) is the convex optimization problem because the second derivatives of $G(n)$ with respect to n is negative. We can use the well–known Lagrangian multiplier method to find the optimal n_{opt}. Accordingly, n_{opt} can be given by[1]

$$n_{opt} \triangleq \arg\min_{n\in[1,N]} \mid G'(n) \mid \tag{27}$$

where $G'(\cdot) \triangleq \frac{\partial}{\partial n} G(n)$.

Equating $G'(n)$ to zero, we can represent n_{opt} in closed–form for a given N as

$$n_{opt} = \arg\min_{n\in[1,N]} \left| \left(N\alpha+1\right) n^{-1} - \alpha\left(\log n + \omega + 1\right) \right|. \tag{28}$$

It can be found from (28) that, in the case when $M=1$, n_{opt} depends on N, and α. Particularly, the argument of $|\cdot|$ in (28) is a decreasing function of n as its first derivative in terms of n is negative in practice. This reveals that for a given N, *the optimal* n_{opt} *satisfying* (28) *can be less than the maximum number* N so that $\bar{R}$ is maximized.

VI. Asymptotic analysis

In this section, we consider an extreme case when the total number N of devices is very large. We study how the optimal number of active devices and the average system rate behave in the extreme case.

[1]For the simplicity in analysis but without loss of generality, we assume that $G(\cdot)$ is continuous and is differentiable at all values of n. But, in practice, n is an integer and thereby the optimal n results in an integer nearest to n_{opt} towards zero.

A. Impact of large N on n_{opt}

As N grows very large (i.e., $N \to \infty$), the behavior of n_{opt} in (28) is investigated. First, let n be $n = N\tau$ for a fixed $\tau \in [1/N, 1]$ such that n ($\leq N$) remains an integer. So, as N increases, n also increases at the fixed rate $\tau = n/N$.

When $N \to \infty$, then the power gain $G(n)$ in (21) is for a given $P_0 (i.e., P_T = P_0)$

$$\lim_{N\to\infty} G(n) = \lim_{N\to\infty} \left(N(1-\tau)\alpha + 1\right) \frac{\log\left(\tau N\right) + \omega}{\log\left(N\right) + \omega} \tag{29}$$

$$\approx \begin{cases} \frac{\omega\, N(1-\tau)\alpha}{\log N} & \text{if } \tau \to \frac{1}{N} \\ 1 & \text{else if } \tau \to 1 \\ \frac{(N(1-\tau)\alpha+1)(\log(\tau N)+\omega)}{\log N} & \text{else } \tau \in (\frac{1}{N}, 1). \end{cases} \tag{30}$$

We can asymptotically observe from this equation that, as N grows for a given P_0, $G(n)$ scales with $N/\log N$ when τ approaches $1/N$; becomes one when τ approaches one; and can be maximized with a proper selection of τ over its intermediate range $\tau \in (1/N, 1)$. Therefore, we can represent that, for large N, n_{opt} is approximated as

$$n_{opt} = \arg\max_{n=\tau N} (30), \quad \text{for} \quad \tau \in \{1/N, ..., 1\}. \tag{31}$$

B. Impact of large N on $\bar{R}$

Similarly, these asymptotic results can be observed in terms of $\bar{R}$. Particularly by inserting (30) into (21), we can observe that for large N, $\bar{R}$ behaves asymptotically as

$$\begin{aligned} \bar{R} &= \frac{1}{N} \sum_{l} \log_2\left(1 + \rho_l (1-\tau) P_0 \ \log N\right) \\ &= \frac{1}{N} \sum_{l} \Theta\left(\log_2 \log N + \log_2\left(\rho_l (1-\tau) P_0\right)\right) \end{aligned} \tag{32}$$

where $\Theta(\cdot)$ denotes the big Theta notation in mathematics. It can be shown from (32) that for large N, $\bar{R}$ increases inversely with respect to τ. That is, as N grows large, $\bar{R}$ is maximized by using the smallest candidate τ (i.e., $n_{opt} = 1$).

From (29)–(32), we present asymptotic summary remarks. For large N, the *round robin transmission is the optimum in order to maximize $\bar{R}$ of the power-limited multi-device scheduling system*. The intuition is that for very large N, the convention scheme makes all the devices waste most of their power for the training towards a small possibility of the data transmission. Such

training power usage can be potentially used for the data. Thus, to avoid wasting their power, the proposed scheme suggests the round robin scheduling for large N. Interestingly, this behavior is opposite to the conventional opportunistic transmission, which is under the fixed power-loading and no selection in scheduling.

Moreover, it is worth pointing out that as per (30)–(32), n_{opt} is asymptotically shown to rely on more P_0, rather than ρ_l. The former is related to the power requirement at devices level, while the latter to the heterogeneous propagation channels.

VII. Simulations and Numerical results

We illustrate simulations and numerical results for the performance of the proposed selective scheduling scheme. For illustrations, let $N \in [2, 50]$ and $n \in \{1, \cdots, 30\}$ are used, along with the training and data power ratio $\alpha = 3/4$ in [26]. For comparison, simulations of the conventional schemes such as the greedy and the PF scheduling are performed over both homogeneous and heterogeneous channels.

A. The cdf performance

In Fig. 3, the cdf of the proposed scheduling scheme with a fixed $n = 8$ is depicted when $N = 10, \alpha = 3/4$ and $\sigma = 1$ dB are used. For comparison, the conventional schemes of both the greedy and the PF scheduling are depicted at the same number $N = 10$ of devices available under the total power requirement. As illustrated in this figure, the greedy scheme slightly outperforms the PF scheme when $\sigma = 1$ dB. Interestingly, Fig. 3 clearly depicts that the proposed scheme can significantly outperform the well-known greedy scheme. For example, to achieve the cdf of 0.3 (or, equivalently, the outage probability of 0.3), it can be shown in Fig. 3 that the proposed scheme obtains the power gain of 4.5 dB over the greedy scheme. The intuition is that the proposed scheme exploits the trade-off between the power-loading balance and the flexible selection scheduling gain while the greedy one is known as the optimum with only the full-selection scheduling.

In Fig. 4, when randomly selecting $n = 6$ devices in scheduling, the cdf of the proposed scheme is depicted when $N = 10, \alpha = 3/4$ and $\sigma = 1$ dB are used. For comparison, the conventional schemes of both the greedy and the PF scheduling are also depicted at the same number $N = 10$ of devices available under the total power requirement. Similar to Fig. 3, Fig. 4

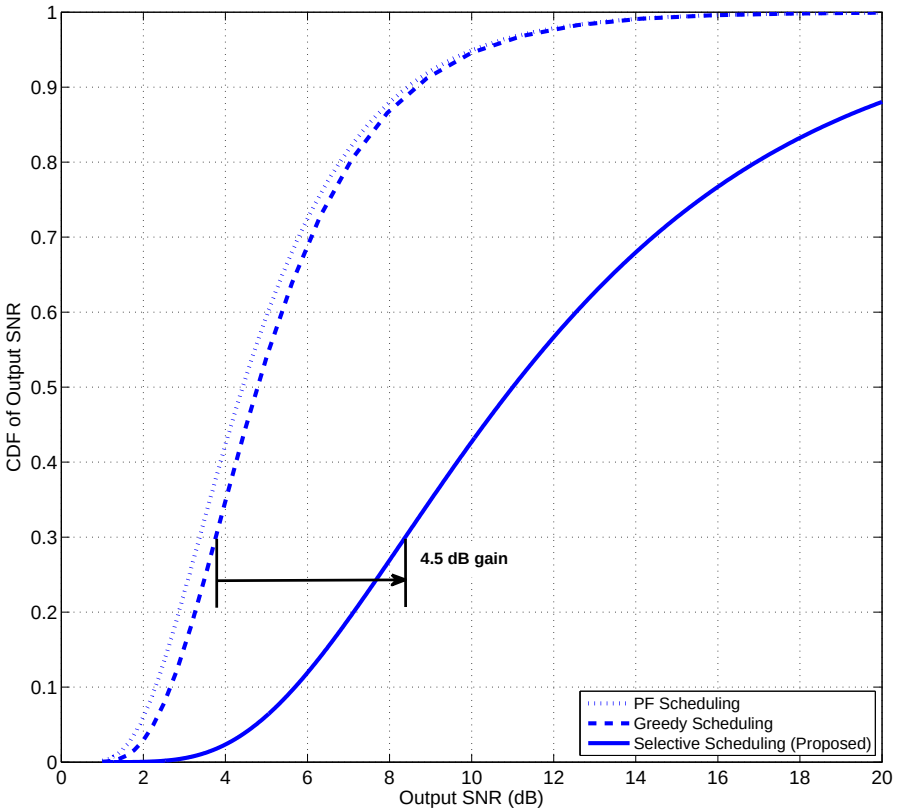

Fig. 3. Comparison of the cumulative distribution functions of the output SNR for the proposed selective scheduling over the conventional schemes (such as the greedy and the proportional fair scheduling). For illustrations, we consider the heterogeneous propagation channels when $N = 10, n = 8, \alpha = 3/4,$ and $\sigma = 1$ dB.

depicts that the proposed scheme can outperform both the conventional schemes (i.e., the greedy and the PF scheduling), with the larger power gain. For example, when having $n = 6$ over $\sigma = 1$ dB moderate heterogeneity in scheduling, to achieve the cdf of 0.3 (or, equivalently, the outage probability of 0.3), Fig. 4 shows that the achievable power gain by the proposed scheme is 9.5 dB over the greedy scheme. From the illustrations, we can observe that smaller n, less the outage probability of the system in scheduling.

Considering $\sigma = 3$ dB for higher heterogeneity in scheduling, Fig. 5 depicts the impact of n on the cdfs of the proposed scheme. For comparison, the cdfs of the conventional schemes are illustrated. Let $N = 10, n \in \{6, 8, 9\}, \alpha = 3/4$ be used. Compared to Fig. 4, it can be seen

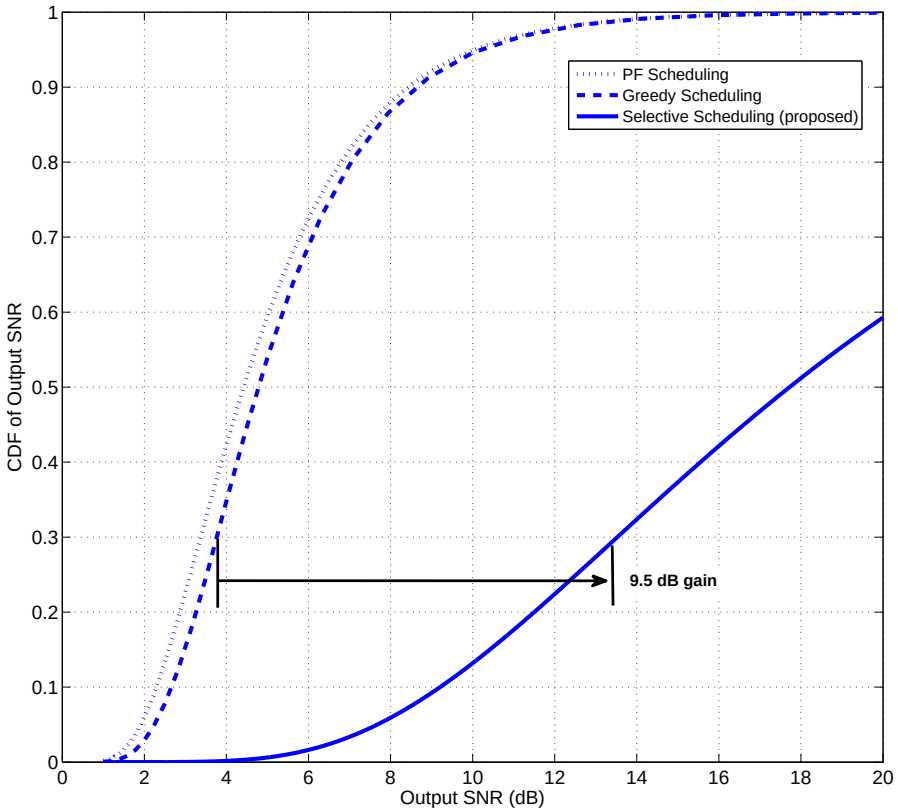

Fig. 4. Comparison of the cumulative distribution functions of the output SNR for the proposed selective scheduling over the conventional schemes (such as the greedy and the proportional fair scheduling). For illustrations, we consider the heterogeneous propagation channels when $N = 10, n = 6, \alpha = 3/4,$ and $\sigma = 1$ dB.

from Fig. 5 that the superiority of the greedy scheme to the PF scheme is larger over higher heterogeneous channels of $\sigma = 3$ dB. Moreover, as for the impact of n at $\sigma = 3$ dB, this figure depicts that for a given threshold y, smaller n results in less the outage probability of the proposed scheme. Therefore, it can be observed from Fig. 5 that the proposed scheme still outperforms both the greedy and the PF schemes, properly selecting the value for n even over highly heterogeneous channels.

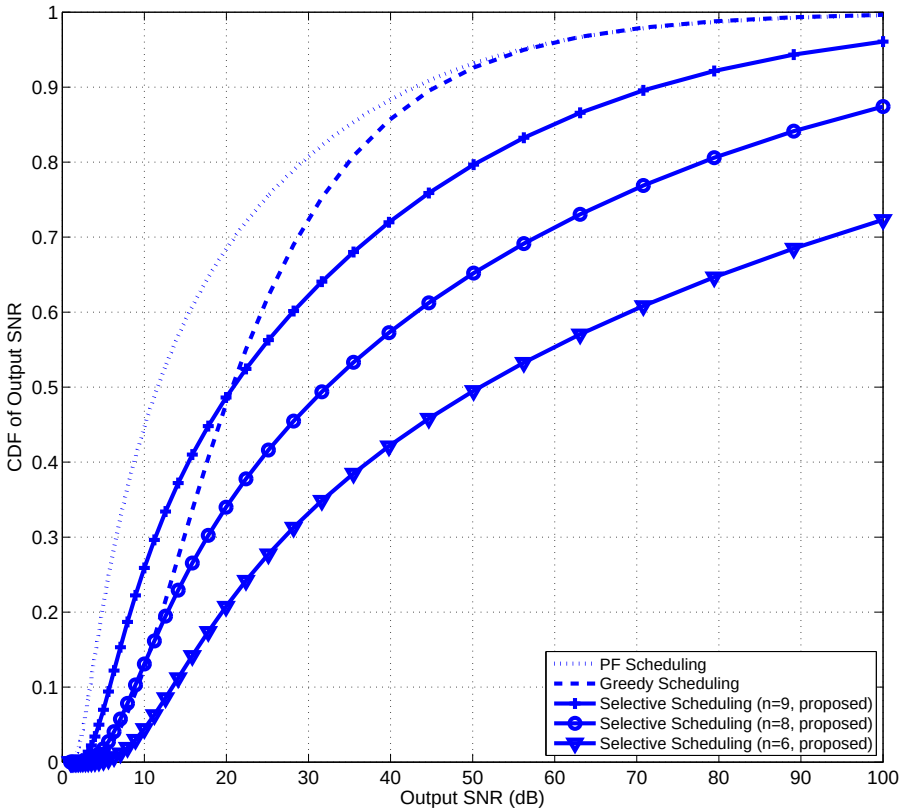

Fig. 5. Impact of the number n of selectively scheduling devices on the cdf. For comparison, the cdfs of the conventional schemes (i.e., the greedy and the proportional fair scheduling) are depicted. We use when $N = 10, n \in \{6, 8, 9\}, \alpha = 3/4$, and $\sigma = 3$ dB.

B. The average system rate performance

In Fig. 6, with several values of $n(\leq N)$, the average system rate $\bar{R}$ is depicted as an increasing function of the number (N) of devices available. For the illustrations, we consider the proposed scheme that exploits the approximate expression for $\bar{R}$ in (20). To validate this approximation, simulations are also depicted in Fig. 6 when $n \in \{4, 8, 10\}, N \in \{10, \cdots, 50\}, \alpha = 3/4$ and homogeneous channels with $\rho = 10$ dB. As seen in this figure, the accuracy between the numerical and simulation results is verified to be good within 0.01 bps/Hz/cell for a wide range of N's.

Moreover, as seen in Fig. 6, there exist turning points of N's beyond which $\bar{R}$ with larger

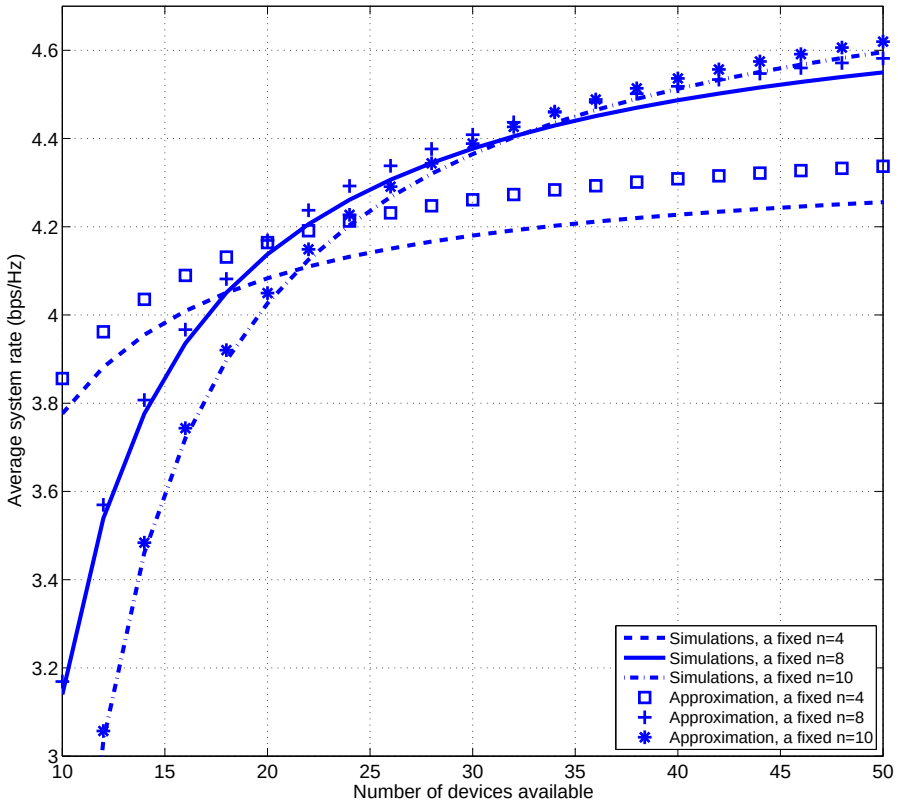

Fig. 6. Average system rate $\bar{R}$ of the proposed scheme has been illustrated with respect to N at several values for n. We use when $N \in \{10, \cdots, 50\}, n \in \{4, 8\}, \alpha = 3/4,$ and homogeneous channels with $\rho_l = 10$ dB for all l. To validate the accuracy of the theoretical results, the simulations are depicted only with markers and no lines while the theoretical ones only with lines and no markers.

n outperforms one with smaller n. For example, it can be shown in Fig. 6 that when $N \geq 18$, selecting $n = 8$ results in $\bar{R}$ higher than ones with $n \in \{4, 10\}$. This validates the analytical results in Section IV.

To further express the impact of n, Fig. 7 now depicts $\bar{R}$ as a concave function of n, for given $N \in \{10, 20, 30\}$. In this illustration, we use when $\alpha = 3/4, \sigma = 3$ dB, and $n \in [0, 100]$ (in percentage) for a given N. As verified in our optimization problem, this figure shows that there exists the optimum value for n to maximize $\bar{R}$. For example, when $N = 20$, it is shown in this figure that the maximum $\bar{R} = 6.933$ is obtained by selecting $n = 6$ (i,.e., 30 per-cent). Also,

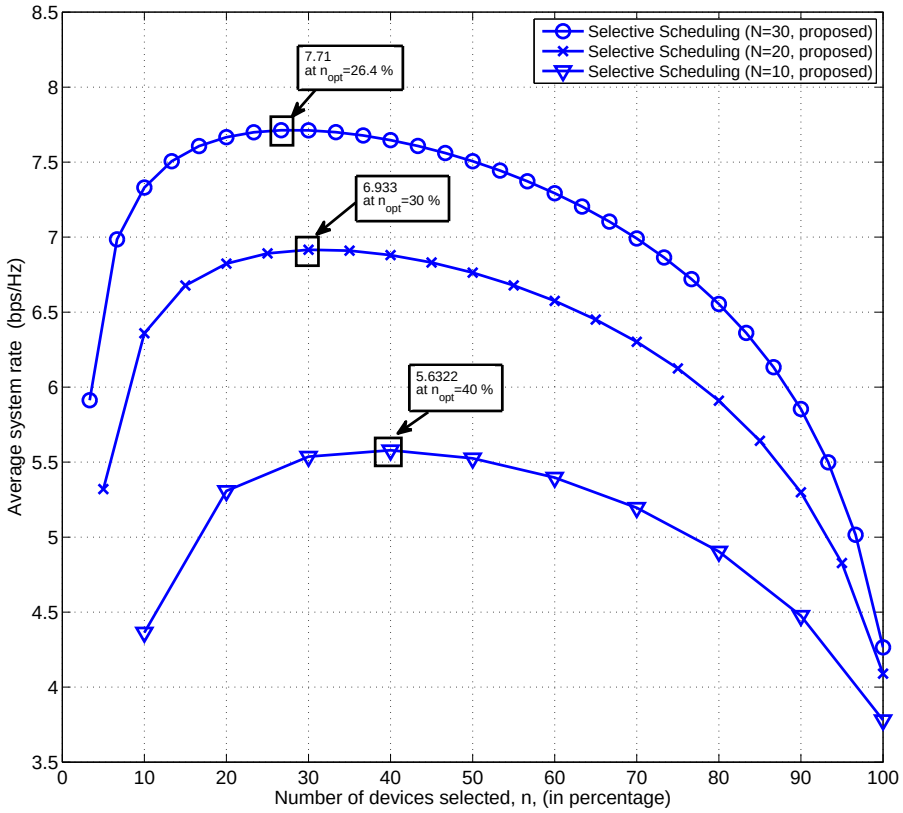

Fig. 7. Average system rate $\bar{R}$ has been illustrated with respect to n at several values for N. For curves, we use when $N \in \{10, 20, 30\}, \alpha = 3/4$, and $\sigma = 3$ dB. For a given N, the optimal n maximizing $\bar{R}$ can be found in this figure.

it is worth mentioning from Fig. 7 that the rate of such an optimum n to N, maximizing $\bar{R}$, decreases as N grows.

As for the impact of N under the power requirement, Fig. 8 depicts $\bar{R}$ as a monotonically increasing function of N. For illustration in this figure, we now normalize the total power P_T at the devices level, regardless of N, and thus P_T remains the same even as N grows. This aims to depict how a large deployment of N devices is allowed to increase $\bar{R}$ of the power-limited multi-devices system. In this figure, we use that $N \in \{2, \cdots, 50\}, n \in \{1, 2, 4, 8, 30\}, \alpha = 3/4$, and $\sigma = 3$ dB. For comparison, we also depict the conventional PF scheme, having a fixed $n = N$. Interestingly, Fig. 8 clearly depicts that the proposed scheme enables $\bar{R}$ to increase in log-scale with N while the PF scheme suffers from decreasing $\bar{R}$. Intuitively, this is because, for large

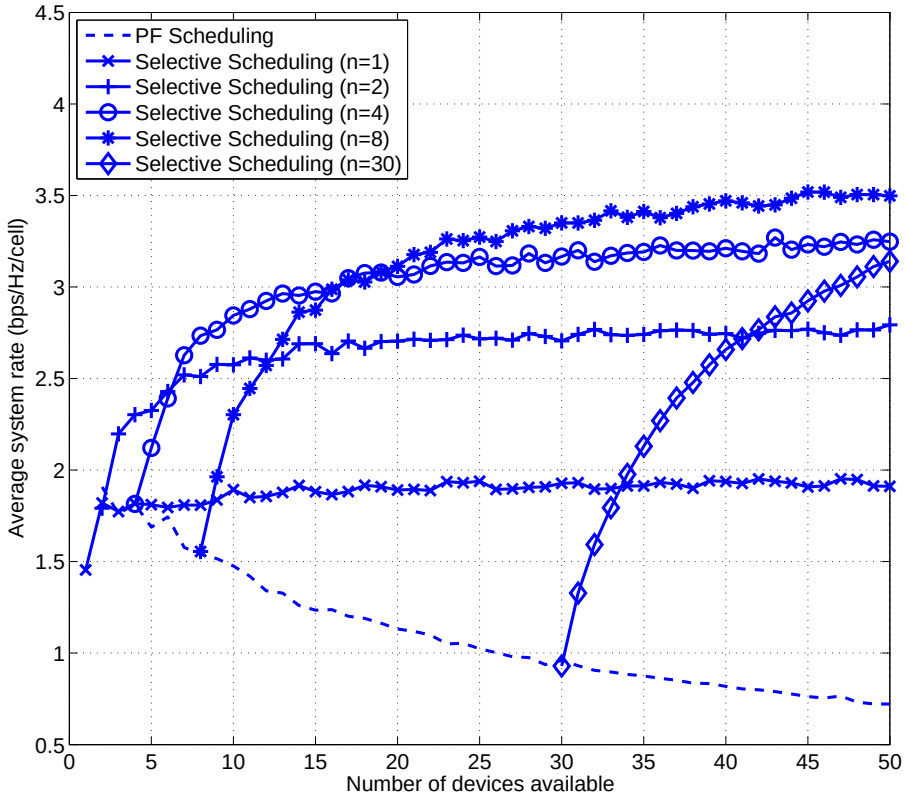

Fig. 8. Average system rate $\bar{R}$ of the proposed scheme has been depicted with respect to N in the case when P_T is normalized, regardless of all N's. For comparison, the conventional proportional fair scheduling is also depicted. In this figure, we use that $N \in \{2, \cdots, 50\}, n \in \{1, 2, 4, 8, 30\}, \alpha = 3/4$ and $\sigma = 3\ dB$.

N, all the devices in the PF scheme wastes their power mostly in the training while n devices randomly selected in the proposed scheduling spend their power, through the power-loading balance and the flexible selection scheduling, to outperform the PF scheme, which validates the analysis in Section VI.

VIII. Conclusions

We studied the multiuser selection scheduling scheme in the non-identically distributed Rayleigh fading channels. We firstly proposed the new proportional fair multiuser selection scheduling scheme suitable to the power-limited multi-device systems. To evaluate the performance of the proposed scheme, we analyzed both the cumulative distribution functions and the average system

rate, deriving their closed-form expressions over the heterogeneous Rayleigh fading channels. To improve the performance with the reduced feedback requirements, we developed the transmit power-loading balance and the partial devices selection scheduling criterion. It can be inferred from this work that, for the power-limited multi-device systems,

- Outage probability performance improves, properly reducing a subset of active devices in scheduling over heterogeneous fading channels.
- Higher average system rate is achievable by jointly designing the number of the partial, active devices and exploiting the power-loading balance.
- Optimum selection scheduling in the sense of the maximum average system rate behaves towards the round–robin scheduling, for very large number of devices. This is opposite to the conventional greedy scheme, which is the optimum with the full feedback requirement and no selection in scheduling.
- Higher the heterogeneous fading, larger the MUDiv gain imposed by the multi-device selection scheduling is.

Based on the outcomes, it is clearly recommended that the proposed scheduling scheme is suitable to the power-limited multi-device systems, especially over the heterogeneous fading.

APPENDIX

A. Proof of the equal probability of accessing the channel

Denote by $\Pr[l = l^*]$ the probability that device $l, \forall l$ are chosen as the best and access the channel. Consider a single antenna $M = 1$ for simple analysis without loss of generality. $\Pr[l = l^*]$ can be written, for given l^* and Σ_D, as

$$\Pr[l = l^*] = \Pr\left[|g_l|^2 = |g_{l^*}|^2, l \in \Sigma_D\right] \tag{33}$$

$$= \Pr\left[|g_l|^2 = |g_{l^*}|^2 \mid l \in \Sigma_D\right] \Pr[l \in \Sigma_D] \tag{34}$$

$$= \frac{n}{N} \prod_{(j)\in\Sigma_D, l\neq(j)} \Pr\left[|g_l|^2 \geq |g_{(j)}|^2\right] \tag{35}$$

where n/N in (35) is based on the system model. Referring to the complementary cdf of $|g_l|^2$ and the moment generating function of the $|g_{(j)}|^2$, we have

$$\begin{aligned}\Pr[l = l^*] &= \frac{n}{N} \prod_{(j)\in\Sigma_D,(j)\neq l}^{n} \mathcal{M}_z\{-1\} \\ &= \frac{n}{N}\frac{1}{4^{n-1}}, \end{aligned} \tag{36}$$

where $\mathcal{M}_z(\cdot)$ is the moment generating function of z and $z = |g_{(j)}|^2$. Therefore, it can be clearly observed from (36) that with the PF scheme, the probability that device l is allowed to access the channel relies only on N and n, being identical for all devices.

REFERENCES

[1] K. David, S. Dixit, and N. Jefferies, "2020 vision the wireless world research forum looks to the future," *IEEE Veh. Technol. Mag.*, vol. 5, pp. 22–29, Sep. 2010.

[2] R. Knopp and P. A. Humblet, "Information capacity and power control in single-cell multiuser communications," in *Proc. of Intl. Conf. on Commun.*, vol. 1, Jun. 1995, pp. 331 – 335.

[3] D. Tse and P. Viswanath, *Fundamentals of wireless communication*, 1st ed. Cambridge University Press, 2005.

[4] P. Svedman, S. K. Wilson, L. J. Cimini, and B. Ottersten, "Opportunistic beamforming and scheduling for OFDMA systems," *IEEE Trans. Commun.*, vol. 55, pp. 941–952, May 2007.

[5] M. Shaqfeh, H. Alnuweiri, and M. S. Alouini, "Multiuser switched diversity scheduling schemes," *IEEE Trans. Commun.*, vol. 60, pp. 2499–2510, Sep. 2012.

[6] Y. L. Che, R. Zhang, and Y. Gong, "On design of opportunistic spectrum access in the presence of reactive primary users," *IEEE Trans. on Communications*, vol. 61, no. 7, pp. 2678–2691, July 2013.

[7] H. Lee, Y. Ko, S. Vahid, and K. Moessner, "Practical spectrum aggregation for secondary networks with imperfect sensing," *IEEE Trans. on Vehicular Technology*, vol. PP, no. 99, pp. 1–1, 2015.

[8] N. Prasad, H. Zhang, H. Zhu, and S. Rangarajan, "Multiuser scheduling in the 3GPP LTE cellular uplink," *IEEE Trans. on Mobile Computing*, vol. 13, no. 1, pp. 130–145, Jan 2014.

[9] R. Morsi, D. S. Michalopoulos, and R. Schober, "Multiuser scheduling schemes for simultaneous wireless information and power transfer over fading channels," *IEEE Trans. on Wireless Communications*, vol. 14, no. 4, pp. 1967–1982, April 2015.

[10] Y. Ko and K. Moessner, "Maximum outage capacity in dense indoor femtocell networks with joint energy and spectrum utilization," *IEEE Trans. Wireless Commun.*, vol. 11, no. 12, pp. 4416–4425, Dec. 2012.

[11] C. Xiong, G. Li, S. Zhang, Y. Chen, and S. Xu, "Energy-efficient resource allocation in OFDMA networks," *IEEE Trans. Commun.*, vol. 60, no. 12, pp. 3767–3778, Dec. 2012.

[12] X. Xiang, C. Lin, X. Chen, and X. Shen, "Toward optimal admission control and resource allocation for LTE-A femtocell uplink," *IEEE Trans. Veh. Technol.*, no. 99, pp. 1–15, Aug. 2014.

[13] Q. Ma and C. Tepedelenlioğlu, "Practical multiuser diversity with outdated channel feedback," *IEEE Trans. Veh. Technol.*, vol. 54, pp. 1334–1345, Jul. 2005.

[14] L. Yang, M. Kang, and M. S. Alouini, "On the capacity–fairness tradeoff in multiuser diversity systems," *IEEE Trans. Veh. Technol.*, vol. 56, pp. 1901–1907, Jul. 2007.

[15] Y. Liu and L. Qiu, "Higher throughput and less feedback: multi-antenna broadcast channels with partial users and limited bits feedback," in *Global Telecommunications Conference, IEEE*, Nov. 2007, pp. 3629–3633.

[16] S. Nai, T. Q. S. Quek, and M. Debbah, "Shadowing time-scale admission and power control for small cell networks," in *Proc. IEEE the 15th Intnal. Symp. on Wireless Personal Multimedia Communications*, Sep. 2012, pp. 628–632.

[17] B. Holter, M.-S. Alouini, G. Oien, and H.-C. Yang, "Multiuser switched diversity transmission," in *Proc. IEEE Vehicular Technology Conf.*, vol. 3, Sep. 2004, pp. 2038–2043.

[18] H. Nam and M. S. Alouini, "Multiuser switched diversity scheduling systems with per-user threshold," *IEEE Trans. Commun.*, vol. 58, no. 5, pp. 1321–1326, May 2010.

[19] Y. Ko, S. A. Vorobyov, and M. Ardakani, "How much multiuser diversity is required for energy-limited multiuser systems?" *IEEE Trans. Signal Process.*, vol. 58, no. 8, pp. 4367–4378, Aug. 2010.

[20] A. Rajanna and N. Jindal, "Multiuser diversity in downlink channels: when does the feedback cost outweigh the spectral efficiency benefit?" *IEEE Trans. Wireless Commun.*, vol. 11, pp. 408–418, Jan. 2012.

[21] J. G. Proakis, *Digital Communications, 4th Ed.* McGraw–Hill, 2001.

[22] H. A. David, *Order Statistics*, 2nd ed. New York: Wiley, 1981.

[23] B. M. Hochwald, T. L. Marzetta, and V. Tarokh, "Multiple-antenna channel hardening and its implications for rate feedback and scheduling," *IEEE Trans. Inf. Theory*, vol. 50, pp. 1893 – 1909, 2004.

[24] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*. Academic Press:San Diego, CA, 7th Ed., 2007.

[25] E. Castillo, *Extreme Value Theory in Engineering*. San Diego: Academic Press, Inc, 1988.

[26] V. K. V. Gottumukkala and H. Minn, "Capacity analysis and pilot-data power allocation for MIMO-OFDM with transmitter and receiver IQ inbalances and residual carrier frequency offset," *IEEE Trans. Veh. Technol.*, vol. 61, no. 2, pp. 553–565, Feb. 2012.